Appendix B

Synopsis of Decision Procedure

The following is a brief synopsis of the Decision Procedure used to evaluate the data from the Dioxin/Furan Tier I Field Study. The complete Decision Procedure is an attachment to the sampling and analysis plan (SAP, Appendix C). The Decision Procedure was used to evaluate the chemical residue and H4IIE-luc bioassay data on polychlorinated dibenzo-p-dioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs) in samples of wildlife tissues to answer the following question:

Are concentrations of PCDD/Fs in representative biota samples collected at RMA greater than those in comparable samples from off-post reference sites?

The first step of the Decision Procedure is to assess the acceptability of the data for statistical analyses. Quality assurance and quality control processes are outlined in the SAP and the laboratory quality control program for the each laboratory. The Decision Procedure indicates that U.S. Army and U.S. Environmental Protection Agency (EPA) personnel audit the laboratories and that a Biological Advisory Subcommittee (BAS) workgroup review the data for general data quality and usability.

The Decision Procedure specifies how concentrations of PCDD/Fs in biota at Rocky Mountain Arsenal (RMA) were statistically compared to those in the same species at off-post reference sites. Three different statistical comparisons of PCDD/F concentrations were made between groups of biota from RMA and off-post reference locations. The first two comparisons examine differences between Toxicity Equivalents (TEQs) and 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD)-EQs in biota from RMA and off-post reference sites, and the third comparison evaluates congener patterns.

In the TEQ method, concentrations of each Ah-R active PCDD and PCDF congener (7 dioxins and 10 furans) present in extracts of each sample are quantified by gas chromatography with high resolution mass spectrometry (HR-GC/MS). The concentration of each congener is then multiplied by a toxic equivalency factor (TEF), which is a fraction that represents the relative potency of a congener as compared to 2,3,7,8-TCDD. The equivalent concentration of 2,3,7,8-TCDD of each congener equals the product of the congener's concentration and its species' TEF. The products are summed over all congeners, and the sum for the sample mixture is denoted as the TEQ concentration. Advantages of the TEQ method are the ability to potentially identify and quantitate each of the 17 Ah-R active PCDD/F congeners in samples, whereby results can be used to examine patterns of contamination and relative contributions of congeners to the TEQ. Disadvantages include quantitative uncertainties due to non-detectable congeners at trace concentrations or when chemical interferences occur. Rarely are all 17 PCDD/F congeners present in samples above method detection limits (MDLs). This problem with inaccuracy and imprecision at trace levels is compounded by quantitative uncertainties in TEFs.

In the bioassay method, extracts from each sample are tested for their ability to induce luciferase activity. Induction of luciferase is closely parallel to biological effects of 2,3,7,8-TCDD, such as the initiation of cellular responses and toxicity (Sanderson et al. 1996). Dose-response curves for the tested extract and for 2,3,7,8-TCDD standards are compared to yield a measure of their relative activity. The result is expressed as the equivalent quantity of 2,3,7,8-TCDD that would have the same enzyme inducing potency as the mixture of PCDD/Fs found in the sample

(Sanderson et al. 1996). Advantages of this method are that the actual biological activity in a sample can be measured in culture and thus estimated *in vivo*. This may partially account for any interactions that would not be apparent in TEQs. Also, there is biological relevance via the Ah-R binding mechanism that can lead to toxicity. This bioassay has the disadvantage that the result can be influenced by chemicals other than PCDD/Fs, and responses cannot always be attributed entirely to PCDD/Fs.

The TEQ and TCDD-EQ yield complementary information, but do not measure exactly the same thing. Rather, they are both estimates of the toxic potency of the mixture of PCDDs and PCDFs in the sample. Both methods were used because of their mutual advantages for increased scientific information.

Although both TEQs and TCDD-EQs provide useful and relevant aggregate measures of the total quantity of PCDDs and PCDFs in the samples, neither utilizes all the information provided by the suite of measurements of congener concentrations. Accordingly, a third method of statistical analysis was used to investigate patterns of relative congener concentrations. This analysis will help identify which congeners are possibly present at elevated concentrations in RMA biota, and may suggest the spatial pattern of contamination at RMA.

To answer the general question posed for the Tier 1 Field Study, greater weight was placed on calculated TEQs, as these measurements represent definitive chemical analyses and are linked to a wider range of environmental and effects data, and because the bioassay does not specifically measure dioxins and furans. There is also a greater regulatory history and acceptance of TEQs for risk assessment than for TCDD-EQ.

The specific statistical tests and area comparisons conducted were dependent upon the species and the character of the data derived from the chemical and bioassay analyses. If the data from samples in RMA and off-post reference groups met the requirements for parametric tests, such as normality of the data distribution and homogeneity of variance, concentrations were compared using standard parametric statistical tests. If data did not meet requirements for parametric statistical tests, their non-parametric equivalents were applied.

Analyses of the patterns of relative concentrations (frequency and magnitude) of congeners included multivariate statistics, such as Principal Components Analysis (PCA) or profile analysis.

The three approaches for statistical comparisons of the data that are outlined above will provide answers to specific Tier 1 Field Study questions that can be formulated as null and alternative hypotheses. The criteria for acceptance or rejection of these testable hypotheses specify a significance of probabilities for Type I error (α) to be less than (<) 0.05 (providing confidence as [1-α] greater than [>] 95 percent) and probability for Type II error (β) to be < 0.20 (producing power as [1-β] > 80 percent).

Because β depends on four main factors (specified α, available sample size, sample variance, and the selected relative effects distance), a relative effects distance has been selected as the greater of either 15 parts per trillion (ppt) TEQ or 50 percent TEQ above the mean concentration found in comparable off-post reference samples. The reason for the lower limit of 15 ppt TEQ

for relative differences between group sample means is that differences less than 15 ppt are meaningless in terms of both biological response and analytical accuracy and precision. Unless noted otherwise in the hypotheses stated throughout this document, the above statistical criteria were applied. However, strict adherence to these requirements did not preclude sound professional observations about the data, such as trends or tendencies with slightly lower levels of statistical significance (ρ less than or equal to 0.1).

Hypotheses (H) Stated to Compare Calculated TEQs

$H1_o$: Mean concentrations of TEQs, based on 2,3,7,8- substituted CDD/Fs, are **not greater** in samples from the RMA on-post (O) population when compared to samples in the same species from an off-post reference (R) population. [H_o: $\mu_o = \mu_r$]

$H1_a$: Mean concentrations of TEQs, based on 2,3,7,8- substituted PCDD/Fs, are **greater** in samples from the RMA on-post (O) population when compared to samples in the same species from an off-post reference (R) population. [H_a: $\mu_o > \mu_r$]

Note: Null hypotheses will be tested separately for American kestrel eggs, great horned owl livers, and carp eggs. Testing of subgroups of kestrel eggs from different locations will be performed according to Hypothesis 2 below.

$H2_o$: Mean concentrations of TEQs, based on 2,3,7,8-substituted PCDD/Fs, are **not greater** in samples from the core (C) and the periphery (P) populations and from off-post reference (R) population. [H_o: $\mu_c = \mu_p = \mu_r$]

$H2_a$: Mean concentrations of TEQs, based on 2,3,7,8- substituted CDD/Fs, are **greater** in samples from either the core (C) or the periphery (P) population compared to the off-post reference (R) population; or are greater in the core (C) population when compared to the periphery (P) population. [H_a: $\mu_c > \mu_r$ or $\mu_p > \mu_r$ or $\mu_c > \mu_p$]

Note: This hypothesis will be tested separately only for American kestrel eggs. The core population area (C) can possibly be defined in the following three ways: (i) conventionally per the USFWS Biomonitoring Plan consisting of birds that potentially nest or feed in RMA Sections 1, 2, 25, 26, 35 and 36 (12 nest box locations designated NW02, NW06, NW07, NW11, NW12, NW25, NW26, NW30t, NW31, SE35, NE35, and NW35); or (ii) selected according to historic findings of elevated levels of dieldrin in kestrel eggs (9 nest box locations designated NW02, NW06, NW25, NW26, NW27, NW31, SE35, NE35, and NW35); or (iii) if dieldrin is analyzed on the same samples as dioxins, the nest box locations that have elevated dieldrin levels ($\geq$ 0.05 parts per million [ppm] in kestrel eggs) will define the core population area.

Hypotheses for Comparing Bioassay TCDD-EQS

$H3_o$: Mean concentrations of TCDD-EQs are **not greater** in samples from the RMA on-post (O) population when compared with samples in the same species from an off-post reference (R) population. [H_o: $\mu_o = \mu_r$]

$H3_a$: Mean concentrations of TCDD-EQs are **greater** in samples from the RMA on-post (O) population when compared with samples in the same species from an off-post reference (R) population. [H_a: $\mu_o > \mu_r$]

Note: This hypothesis will be tested separately for American kestrel eggs, great horned owl livers and carp eggs. Testing of subgroups of kestrel eggs from different locations will be done according to Hypothesis 4 below.

$H4_o$: Mean concentrations of TCDD-EQs are **not greater** in samples from the core (C) and periphery (P) populations and from the off-post reference (R) population. [H_o: $\mu_c = \mu_p = \mu_r$]

$H4_a$: Mean concentrations of TCDD-EQs are **greater** in samples from either the core (C) or periphery (P) population when compared to the off-post reference (R) population; or are greater in the core (C) population when compared with the periphery (P) population.
[H_a: $\mu_c > \mu_r$ or $\mu_p > \mu_r$ or $\mu_c > \mu_p$]

Note: This hypothesis will be tested only for American kestrel eggs. The core population area (C) can possibly be defined in three ways: (i) conventionally per the USFWS Biomonitoring Plan consisting of birds that potentially nest or feed in RMA Sections 1, 2, 25, 26, 35 and 36 (12 nest box locations designated NW02, NW06, NW07, NW11, NW12, NW25, NW26, NW30t, NW31, SE35, NE35, and NW35); or (ii) selected according to historic findings of elevated levels of dieldrin in kestrel eggs (9 nest box locations designated NW02, NW06, NW25, NW26, NW27, NW31, SE35, NE35, and NW35); or (iii) if dieldrin is analyzed concurrently with dioxins, the nest box locations that have elevated dieldrin levels (≥ 0.05 ppm in American kestrel eggs) will define the core population area.

Hypothesis for Comparing Congener Patterns

$H5_o$: Patterns of relative concentrations (ratios of congeners) of PCDD/Fs are **not different** in samples from the RMA on-post (O) population compared to patterns in samples in the same species from an off-post reference (R) population. [H_o: $\mu_o = \mu_r$]

$H5_a$: Patterns of relative concentrations (ratios of congeners) of PCDD/Fs are **different** in samples from the RMA on-post (O) population compared to patterns in samples in the same species from an off-post reference (R) population. [H_o: $\mu_o \neq \mu_r$]

Note: This hypothesis will be tested separately for American kestrel eggs, great horned owl livers, and carp eggs.

Individual decision flowcharts were developed for TEQ, TCDD-EQ, and pattern analysis for each species. Once a decision for each type of analysis for each species was reached, an over-all decision matrix and flowchart were followed to answer the overall question as stated above (**Tables B-1** through **B-4** and **Figure B-1**).

Table B-1. Decision matrix for American kestrel eggs and great horned owl livers to support the evaluation of PCDD/Fs as COCs[1] at the RMA

Step V in column 5 below addresses the general question to be answered by the Biological Assessment Subcommittee (BAS) for this Tier I Field Study, stated as: ***Are concentrations[2] of PCDD/Fs in biota samples from the RMA greater than those in the same species collected from the selected off-post reference locations?***

Step I: Data Usability	Step II: TEQ ($H1_o$ or $H2_o$)	Step III: TCDD-EQ ($H3_o$ or $H4_o$)	Step IV: Pattern Analyses ($H5_o$)	Step V: BAS's Answer for Overall Decision[3, 4]	Examples of the BAS's considerations for professional interpretation of the Overall Decision
Evaluated	Reject H_o	Reject H_o	Reject H_o	YES	Probable COC at RMA.
Evaluated			Accept H_o	YES or Inconclusive	Perform mass-balance[5] with REPs (relative effect potencies).
Evaluated	Reject H_o	Inconclusive	Reject H_o	YES	Probable COC at RMA.
Evaluated			Accept H_o	YES or Inconclusive	Perform mass-balance[5] with REPs.
Evaluated	Reject H_o	Accept H_o	Reject H_o	YES	Possible[6] COC at RMA.
Evaluated			Accept H_o	YES or Inconclusive	Perform mass-balance[5] with REPs.
Evaluated	Accept H_o	Reject H_o	NA	Recalculate TEQs including PCBs	After recalculating the TEQs including PCBs, repeat the statistical analysis, and use the sub-matrix below.
Evaluated	Inconclusive	Reject H_o	NA	Recalculate TEQs including PCBs	After recalculating the TEQs including PCBs, repeat the statistical analysis, and use the sub-matrix below.
Evaluated	Inconclusive	Inconclusive	Reject H_o	Inconclusive	May indicate a small local PCDD/F source.
Evaluated			Accept H_o	NO	Uncertain toxicity equivalent factors (TEFs) and trace analysis may be cause for TEQ.
Evaluated	Inconclusive	Accept H_o	Reject H_o	Inconclusive	May indicate a small local PCDD/F source.
Evaluated			Accept H_o	NO	Uncertain TEFs and trace analysis may be cause for TEQ.
Evaluated	Accept H_o	Inconclusive	Reject H_o	Inconclusive	May indicate a small local PCDD/F source.
Evaluated			Accept H_o	NO	Possible non-PCDD/Fs causing slightly higher bioactivity.
Evaluated	Accept H_o	Accept H_o	Reject H_o	Inconclusive	May indicate a small local PCDD/F source.
Evaluated			Accept H_o	NO	Probably not a COC at RMA.

Table B-2. Decision Sub-matrix for American kestrel eggs and great horned owl livers to evaluate PCB contributions at the RMA for outcomes when the null hypothesis is rejected for Step III TCDD-EQ but accepted or inconclusive for Step II TEQ

Step V in column 5 addresses the general question for this Tier I Field Study: ***Are concentrations[2] of PCDD/F in biota samples from the RMA greater than those in the same species collected from the selected off-post reference locations?***

Recalculate the TEQ including PCBs for Step II, and then use the following matrix for decision for the overall outcome.

Step I: Data Usability	Step II: TEQ ($H1_o$ or $H2_o$)	Step III: TCDD-EQ ($H3_o$ or $H4_o$)	Step IV: Pattern Analysis ($H5_o$)	Step V: Overall Decision[3,4]	Examples of considerations for interpretation of Overall Decision
Evaluated	Reject H_o	Reject H_o	Reject H_o	Inconclusive	May indicate a small local PCDD/F source; however, PCB congeners account of majority of differences.
Evaluated			Accept H_o	NO	This outcome may indicate that PCB congeners are significantly greater for RMA samples than off-post reference samples. The BAS will consider the implications.
Evaluated	Accept H_o	Reject H_o	Reject H_o	Inconclusive	May indicate a small local PCDD/F source.
Evaluated			Accept H_o	NO	Possible other agonist causing bioactivity.

Notes: (for Tables A and B):

1. COC (contaminant of concern) is an EPA term for a chemical that has both a source and a potential for release from a site, as per EPA Guidance (EPA 1989) that is based on Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) and National Contingency Plan (NCP) regulations. The BAS agreed to use a stepwise scientific approach that evaluates the weight and strength of the major "lines of scientific evidence" from tiered biological studies at the RMA, which provide site-specific information to evaluate whether PCDD/Fs may be COCs. Using this stepwise approach to reach the overall decision in Step V above, Step I (not shown) was performed first to ensure the adequacy of data for further valid biostatistical evaluations, and then the BAS considered the anticipated combinations of possible results as shown in Steps II through IV. The possible outcomes in the matrix are sorted in descending order with the strongest evidence for existence of COCs at the top and the strongest evidence for absence of COCs at the bottom, with more weight being given to the results from the TEQ analyses in Step II.
2. Concentration, as used in this context, means "toxic-equivalents" of 2,3,7,8-TCDD that are generated by the 17 PCDD/F congeners with Ah-R agonist activity. It is important to note that only Step II (TEQ) provides results from a direct measure of PCDD/F concentrations, although those measurements can become less certain near the analytical detection limits due to measurement errors and due to uncertainties in TEFs; additionally, Step III (TCDD-EQ) can provide an indirect measure of PCDD/F concentrations, provided that the bioassay results are not overshadowed by other chemicals with Ah-R activity.
3. An "inconclusive" decision indicates that the general question posed cannot be answered as "yes" or "no" with sufficient scientific confidence. An inconclusive outcome will result in further ecotoxicological analysis of the problem by the BAS.
4. The BAS recognizes that bioassay derived TCDD-EQ concentrations might not reflect analytically derived TEQ concentrations because biota extracts may contain substantial amounts of other types of Ah-R agonists or antagonists (e.g., PCBs, polycyclic aromatic hydrocarbons, polychlorinated naphthalenes, etc.). If such other Ah-R agonists or antagonists are present in samples at sufficiently high concentrations, they will likely influence the TCDD-EQ concentrations while not being totally accounted for in the chemical residue analyses. Therefore, while TCDD-EQ results by themselves cannot answer the

general question posed in the Tier 1 Field Study, TCDD-EQs can be used in a weight-of-evidence approach to help guide (a) the interpretation of toxicological significance (especially if PCDD/Fs have the predominance of Ah-R activity), and (b) possible future studies at the RMA. The BAS generally recognizes that TCDD-EQs, if not overshadowed by other Ah-R activity, can potentially show differences (similar to TEQs) in PCDD/F concentrations on- and off-post.

5. This overall answer depends on the results of the pattern analyses: (a) if the Principal Components Analysis (PCA) visual patterns and/or cluster analyses and profile analyses of relative concentrations of PCDD/F congeners are the same, but the masses of PCDD/Fs are substantially greater on-post than in off-post samples, then the outcome is "yes," or (b) if the masses are similar in this event, then the outcome is "inconclusive."

6. The suggested interpretation of the outcome for this scenario is downgraded to "possible COC" from "probable COC," because this situation is anticipated to occur from a small difference between groups with relatively low TEQs that may be barely significant ($p < 0.05$); therefore, there would likely be greater uncertainty in this outcome, since the results may be driven by error in trace-level detection limit concentrations coupled with uncertain TEFs.

Table B-3. Decision Matrix for Combined Results for Terrestrial Species to Support the Evaluation of PCDD/Fs as COCs at the RMA

Column 4 addresses the general question for this Tier 1 Field Study: ***Are concentrations of PCDD/F in biota samples from the RMA greater than those in the same species collected from the selected off-post reference locations?***

Text Reference[1]	American Kestrel Decision	Great Horned Owl Decision	Overall Terrestrial Species Decision
V.B.1	YES	YES	YES
V.B.1	YES	NO	YES
V.B.1	YES	Inconclusive	YES
V.B.1	Inconclusive	YES	YES
V.B.1	NO	YES	YES
V.B.2	NO	Inconclusive	Inconclusive
V.B.2	Inconclusive	NO	Inconclusive
V.B.2	Inconclusive	Inconclusive	Inconclusive
V.B.3	NO	NO	NO

[1] Text references are from BAS (2000). *Rocky Mountain Arsenal Dioxin/Furan Tier I Field Study Sampling and Analysis Plan*.

Table B-4. Decision Matrix for Carp Eggs to Support the Evaluation of PCDD/Fs as COCs at the RMA

Column 5 addresses the general question for this Tier 1 Field Study: ***Are concentrations of PCDD/F in biota samples from the RMA greater than those in the same species collected from the selected off-post reference locations?***

Step I: Data Usability	**Step II: TEQ ($H1_o$)**	**Step III: TCDD-EQ ($H3_o$)**	**Step IV: Pattern Analysis ($H5_o$)**	**Overall Outcome**
Evaluated	Reject $H1_o$	Reject $H3_o$	Use to determine principal components	YES
Evaluated	Reject $H1_o$	Inconclusive	Reject $H5_o$	YES
Evaluated	Inconclusive	Reject $H3_o$	Reject $H5_o$	YES
Evaluated	Reject $H1_o$	Inconclusive	Accept $H5_o$	Inconclusive
Evaluated	Inconclusive	Reject $H3_o$	Accept $H5_o$	Inconclusive
Evaluated	Inconclusive	Inconclusive		Inconclusive

Figure B-1. Flowchart of Overall Decision Procedure for American Kestrel Eggs and Great Horned Owl Livers to Support the Evaluation of PCDD/Fs as COCs at RMA

Are concentrations of PCDD/F in biota samples from RMA greater than those in the same species collected from the selected off-post reference locations?